(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 June 2003 (26.06.2003)

PCT

(10) International Publication Number
WO 03/052627 A1

(51) International Patent Classification[7]: G06F 17/30, 17/20

(21) International Application Number: PCT/AU02/01719

(22) International Filing Date: 18 December 2002 (18.12.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PR 9589 18 December 2001 (18.12.2001) AU

(71) Applicant (*for all designated States except US*): **TELSTRA NEW WAVE PTY LTD** [AU/AU]; ACN 070 562 935, 242 Exhibition Street, MELBOURNE, Victoria 3000 (AU).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): **RYAN, Simon, David** [AU/AU]; 11 Sandgate Avenue, GLEN WAVERLEY, Victoria 3150 (AU). **RASKUTTI, Bhavani** [AU/AU]; 4 Empress Road, SURREY HILLS, Victoria 3127 (AU). **PHIET, Do, Quang** [AU/AU]; 25 Andleigh Drive, MULGRAVE, Victoria 3170 (AU). **SEMBER, Peter, Paul** [AU/AU]; 22 Neville Street, CARNEGIE, Victoria 3163 (AU).

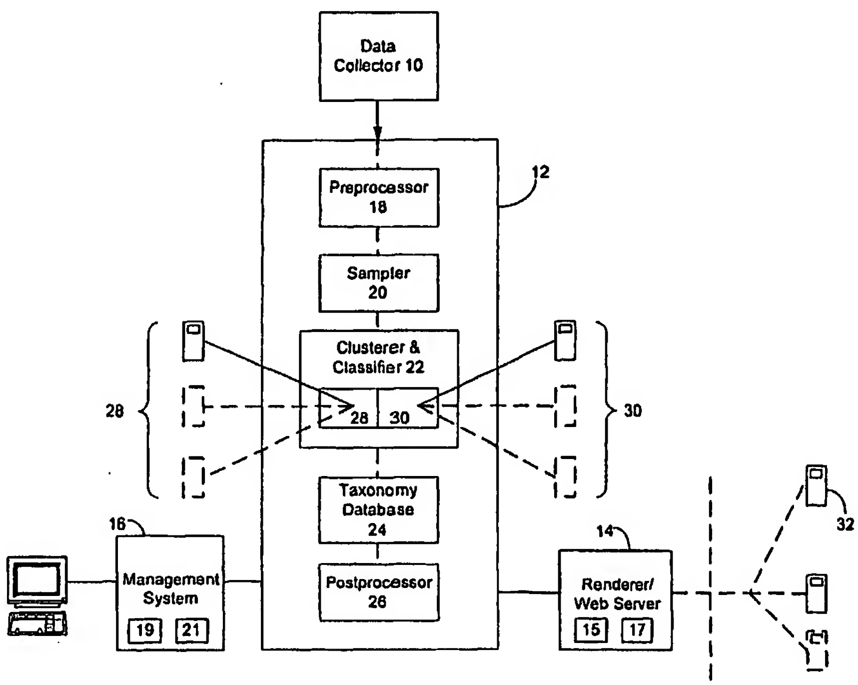
(74) Agent: **DAVIES COLLISON CAVE**; WEBBER, David, Brian, PRYOR, Geoffrey, Charles, LESLIE, Keith, 1 Little Collins Street, MELBOURNE, Victoria 3000 (AU).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,

[Continued on next page]

(54) Title: INFORMATION RESOURCE TAXONOMY



(57) Abstract: An information resource taxonomy system, including a data collector for collecting information resources from a communications network; and a taxonomy generator for generating a taxonomy represented by a hierarchy of resource clusters, using cluster criteria generated from the collected resources. The system includes an editor for editing the criteria, and a renderer for generating linked document data for displaying the hierarchy. A parallel cluster search system is used to evaluate clusters in parallel. The system also includes a parallel classifier for classifying further collected resources.

WO 03/052627 A1

ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

INFORMATION RESOURCE TAXONOMY

FIELD OF THE INVENTION

The present invention relates to taxonomies for information resources, and in particular to a system and process for generating a taxonomy for a plurality of information resources in a communications network.

BACKGROUND

The enormous number of stored electronic documents and other information resources available in modern communications networks such as the Internet poses particular problems for classification and categorisation. For example, the world wide web provides access to an ever-increasing number of electronic documents, many of them generated dynamically, and it is often difficult to retrieve a document of interest without knowing in advance at least part of an identifier, address or locator for the resource. For this reason, search engines have been developed which attempt to generate lists of relevant documents in response to keywords typed in by a user. However, such searches are limited by the choice of keywords entered by the user. As an alternative, directories of web resources have been created by manual vetting and categorisation of web documents into hierarchical category structures known as web directories. These directories are extremely useful for locating relevant documents once a particular category has been chosen. However, the development of these directories is a challenge in itself. For example, companies such as Yahoo! have employed more than 300 people for maintaining the structure of their online directory. This level of expenditure is not justifiable for most companies. More recently, some solutions have appeared which replace the manual vetting with automatic classification based on a manually created taxonomy. Although this alleviates the problem to some extent, the manpower needed to create and maintain the appropriate taxonomy is still considerable. It is desired, therefore, to provide an improved system and process for generating a taxonomy for information resources in a communications network, or at least a useful alternative.

SUMMARY OF THE INVENTION

In accordance with the present invention there is provided a process for generating a taxonomy for a plurality of information resources in a communications network, including:

collecting said resources from said network;

generating cluster criteria from said resources; and

generating said taxonomy as a hierarchy of resource clusters based on said criteria.

The present invention also provides an information resource taxonomy system, including

a data collector for collecting information resources from a communications network; and

a taxonomy generator for generating a taxonomy represented by a hierarchy of resource clusters, using cluster criteria generated from said resources.

BRIEF DESCRIPTION OF THE DRAWINGS

Preferred embodiments of the present invention are hereinafter described, by way of example only, with reference to the accompanying drawings, wherein:

Figure 1 is a schematic diagram of a preferred embodiment of an information resource taxonomy system;

Figure 2 is a flow diagram of a data collection process executed by a data collector of the system;

Figure 3 is a flow diagram of a pre-processing process executed by a pre-processor of the system; and

Figure 4 is a graph of the goodness value of a document set as a function of the cluster threshold.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

As shown in Figure 1, an information resource taxonomy system includes a data collector 10, a data processing system 12, a renderer 14, and a management system 16. The

taxonomy system executes a taxonomy generation process that automatically generates a taxonomy from structured or unstructured documents or other information resources, and can be used to maintain the taxonomy. The taxonomy is a hierarchical tree structure that organizes resources into clusters or nodes based on their similarity, and can include the resources themselves. The taxonomy is subsequently used by the renderer 14 to generate markup code such as HTML, XML, or ASP that provides an interactive, hierarchical view into the space of documents or other information resources. A user of the Internet can view the hierarchy and open individual documents or other information resources over the Internet using a web browser 32 to access the markup code generated by the renderer 14 and generate a graphical display of the hierarchy. The taxonomy system can be applied to a variety of taxonomy generation tasks such as site management of corporate intranets and external web sites.

An administrator of the taxonomy system can login to the system from a terminal associated with the management system 16. The administrator can then submit to the taxonomy system a text file that defines the taxonomy specifications, *i.e.*, the taxonomy creation tasks to be performed by the system. This file includes a list of universal resource indicators (URIs) and a corresponding list of 'include' specifications. The URIs indicate high-level domains that are to be clustered or categorised by the taxonomy system, and the 'include' specifications indicate the types of documents that are to be included in the taxonomy. For example, it may be desired to include only textual documents in one or more of the following formats: HTML, text, Microsoft Word®, FrameMaker, and StarOffice. The text file containing these specifications is sent to the data collector 10.

The components of the taxonomy system can be implemented using standard computer system hardware and adding unique software modules. For example, the data collector 10 and the renderer 4 are 850 MHz Pentium 3 and 1.5 GHz Pentium 4 personal computers, respectively, each running a Linux operating system. The data processing system 12 is a Sun Ultra Enterprise four-CPU server running a Solaris 8 operating system. The management system 16 is a 1.5 GHz Pentium 4 personal computer running a Windows XP operating system. The data processing system 12 includes a number of data processing

modules 18 to 26, including a pre-processor 18, a sampler 20 a clusterer and classifier 22, a taxonomy database 24, and a post processor module 26. The data processing system 12 can further include parallel clusterers 28, and/or parallel classifiers 30. The renderer 14 includes a taxonomy rendering module 15 and a web server module 17. The management system 16 includes a process management component 19 and an editor module 21. Whilst these modules are preferably implemented by software code, at least some of the processing steps executed by the modules, described below, may be implemented by hardware circuits such as application- specific integrated circuits (ASICs).

The data collector 10 executes a data collection process, as shown in Figure 2. The data collection process begins at step 34 when the taxonomy specifications are received. The collector 10 uses the specifications to navigate or "crawl" the Internet at step 36, starting at the top level domains provided by the URI lists and progressing down to sub-domains thereof. The crawling process is known in the art. Briefly, the data collector 10 performs HTTP GET requests to network servers indicated by the provided URIs, or by links within HTML data previously retrieved from the network, including only those links that match the include specifications. For each document retrieved, the data collector 10 converts any documents that are not in HTML into HTML at step 38. The resulting HTML data is then sent to the data processing system 12 at step 40. If the data collector 10 has exhausted all of the hyperlinks contained within documents retrieved from the network, then the process branches at step 42 to return to step 34, and waits for the next category specification to be submitted by an administrator. Alternatively, if it is determined at step 42 that more data needs to be collected, the process branches back to step 36 in order to retrieve more data from the network.

HTML data sent to the data processing system 12 from the data collector 10 is received by the pre-processor 18. Alternatively, HTML data can be directly submitted to the pre-processor 18 by the administrator using the management system 16. The pre-processor 18 executes an HTML processing process, as shown in Figure 3. The process begins when HTML data is received by the pre-processor 18 at step 44. Metadata tags are then extracted from the HTML data at step 46. This is achieved by regular expression matching

on predefined patterns such as the HTML tags <TITLE> <META...> and so on. Meta information is included in the output from the pre-processor 18 as text-delimited additions to the data. The delimiters are text markups that do not normally occur in the data, *e.g.*, "xxxxxxxx:". The remaining data is then processed at step 48 by a filter that removes data that is not considered to be important. This includes removing text that appears likely to be a component of an advertising table or banner. Commonly occurring noise strings are removed by stoplists or by statistical analysis. For example, noise reduction can be achieved by building a frequency table of strings found in the document set. These strings are the characters found between matching pairs of HTML tags, such as <TD> and </TD>. A string is removed from the document set if its occurrence frequency exceeds a threshold value. At step 50, the pre-processor 18 converts the remaining HTML to text by removing HTML tags. The resulting text document is then sent to the sampler 20 at step 52. The sampler 20 samples a fixed fraction of incoming documents, as described below. The sample documents are then processed by the clusterer/classifier 22.

The clusterer 22 partitions the documents based on their content. It does this by forming groups or clusters of documents based on their natural affinity rather than requiring a pre-specified number of categories. The clustering and feature selection processes are based upon processes described in the specification of International Patent Application No. PCT/AU01/00198 ("the TACT specification"), incorporated herein by reference. First, each document is represented by a word frequency vector including words from the document and their frequencies of occurrence, where some words are excluded using feature selection criteria. A numeric similarity measure is then determined as a function of any two word vectors to determine the similarity of any two documents. For example, a new cluster can be formed by two documents if their similarity falls within a threshold similarity value for clustering. Once formed, a cluster is characterised by a word frequency vector that is the average of the word frequency vectors of its constituent documents. This average word frequency vector is referred to as the cluster centroid. The similarity measure used is the cosine similarity function, described in the TACT specification. The clustering process uses this similarity measure to group similar documents into clusters by assigning each document to the most similar cluster. An

optimal similarity threshold value for creating clusters from a given document set is determined by creating different groupings of the documents at different thresholds and then evaluating these to determine the best grouping, as described in *An Evaluation of Criteria for Measuring the Quality of Clusters* by B. Raskutti and C. Leckie, pp. 905-910, in Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence, 1999. This evaluation is based on minimising a goodness value that is based on the similarity of documents within clusters, which tends to reduce the number of documents in each cluster, and the separation of cluster centroids from the global centroid, which encourages larger clusters. For example, a goodness value for a document set can be determined by simply summing these two values.

Hierarchical clustering is achieved by iterative clustering of larger, less coherent clusters. The coherence of a cluster is determined by the intra-cluster similarity value of the cluster. If the documents in a cluster are very similar, *i.e.*, the similarity values of each document with the cluster centroid fall within a similarity threshold for coherence, then the cluster is deemed coherent. If this criterion is not met, then documents within the cluster are formed into sub-clusters of the original cluster. These sub-clusters are sub-nodes of the original parent cluster or node, thus forming a hierarchy of clusters or nodes. By performing this sub-clustering iteratively, a hierarchical tree structure of coherent clusters is formed, to provide the taxonomy. The computational complexity of this clustering process is proportional to n, the number of documents, K, the number of threshold evaluations and m, the average number of clusters per threshold.

The clustering process includes several steps for alleviating some of the scalability issues by reducing n and K. Whilst m is much smaller that n, it is proportional to n, therefore reducing n also reduces m. In one form, execution time is reduced by using percentage-based random sampled clustering of the document space whereby the sampler 20 provides a fixed fraction of the document space to the clusterer 22 for clustering.

A second form is provided by stopping the clustering process after a predefined time interval in order to generate a clustered sample of the document space. These two forms of optimisation can be used independently or in conjunction.

After the initial clustering has been performed on a subset or sample of the document set, the remaining documents are subsequently assigned to the clusters by one of three processes. The first process simply classifies documents into the existing clusters using the existing cluster centroids. That is, a new document is added to an existing cluster if its similarity to the cluster centroid falls within a fixed threshold similarity value. Any documents failing the threshold evaluation criteria for all clusters are set aside for later clustering.

The second process uses the sample document clusters as a training set for an alternative document classification system. In this case, a support vector machine (SVM) is used as an alternative classifier. The SVM is described in the specification of International Patent Application No. PCT/AU01/00415, incorporated herein by reference. As with the first process, any documents not classified are set aside for later clustering.

The third process simply continues to cluster, but using the optimal threshold similarity value determined whilst clustering the initial sample documents. This process forms new clusters for new documents that are not similar to the existing clusters.

Each of these three processes is an approximation and assumes that the original sample is representative of the complete (or future) document space. Consequently, errors are introduced over time as more documents are added to the clusters due to cluster centroid drift. Two processes are used to combat this effect. In the first process, the coherence of the clusters is maintained as the number of documents *n* increases by reducing the similarity threshold with increasing *n*.

In the second process, a new random sample better representing the population is determined as the document collection grows. The new sample is used as a metric for evaluating the optimality of the existing clusters and/or as a means for determining a new quasi-optimal similarity threshold value for subsequent re-clustering of the document space to improve accuracy.

To reduce the time required by the search for an optimal or quasi-optimal similarity threshold value, cluster formation with different threshold values can be performed by different threads or on different processors in an SMP or distributed processing framework, such as the parallel clusterers 28. The time spent searching for the optimum threshold value is also reduced by using an efficient search process based on knowledge of the topography of the goodness vs threshold similarity curve. For example, Figure 4 is a graph of the goodness value of a document set, as described above, as a function of the logarithm of the similarity threshold value for cluster formation. The solid line 54 joining data points has a well defined minimum 56 at a log (threshold) value near 0.2. The general shape of this graph is typical of all document sets. Knowing the approximate shape of this graph allows the optimal threshold value for a particular document set to be located rapidly.

The taxonomy produced by clustering is stored in the taxonomy database 24. After clustering, the postprocessor module 26 augments the clustered data by extracting titles from metadata of each document, and adding summary text generated by the clustering process, as described in the TACT specification. In cases where access logs (*i.e.*, web server or proxy cache logs) are available for each document, the clusters and/or documents within each cluster can be ranked using the access frequency of each document. For example, on a corporate web server, the most popular pages are listed near the top of each category listing, and/or the most popular categories are listed near the top of a listing of categories.

The management system 16 includes an editor 21 that allows the administrator to manually edit a taxonomy to create a new document hierarchy. This new structure can then be used as the training set for adding further documents to the database using the classifier function

of the clusterer/classifier 22. The speed of document classification by the categorisation system can be improved by using the parallel classifiers 30 to classify many documents in parallel.

The editor 21 offers a number of editing functions, including moving branches of the hierarchical taxonomy to other branches, editing meta descriptions for documents and branches, and creating, deleting, and merging new branches in the taxonomy. The editor 12 presents information from the taxonomy database 24 using HTML forms. Changes can then be made to the taxonomy by modifying input fields in the forms and then submitting the changes via submit buttons of the forms.

The taxonomy rendering module 15 of the renderer 14 generates dynamic web pages using the taxonomy database 24 to provide structure to the original resource content. These web pages can be accessed by providing to the web browser 32 a URI associated with the web server module 17. The visual presentation provided by these web pages is derived from a configuration file detailing the arrangement of the various fields on the rendered page. The pages represent a web 'view' into the hierarchy using a 'directory' style wherein the URI of the displayed page corresponds to the position or branch within the taxonomy that is being browsed. Each level in the 'view' can contain documents and/or categories, *i.e.*, deeper branches in the taxonomy. Browsing into a category produces a new view with a greater level of specificity. Each branch in the taxonomy is initially labelled automatically by extracting descriptive information from the data during taxonomy generation, as described above, and is manually editable by invoking the editor module 21 of the management system 16. Documents are presented using their titles and summaries. Browsing to the document opens the document or a representation of the document.

Many modifications will be apparent to those skilled in the art without departing from the scope of the present invention as herein described with reference to the accompanying drawings.

CLAIMS:

1. A process for generating a taxonomy for a plurality of information resources in a communications network, including:
collecting said resources from said network;
generating cluster criteria from said resources; and
generating said taxonomy as a hierarchy of resource clusters based on said criteria.

2. A process as claimed in claim 1, including generating linked document data for displaying said taxonomy.

3. A process as claimed in claim 2, wherein said linked document data includes markup language data.

4. A process as claimed in claim 2, wherein said linked document data includes metadata of said resources.

5. A process as claimed in claim 1, including generating descriptive text for said resources and descriptive text for each node of said hierarchy.

6. A process as claimed in claim 1, wherein said cluster criteria is used to classify said resources.

7. A process as claimed in claim 1, wherein said resources include dynamically generated content of said network.

8. A process as claimed in claim 1, wherein components of said hierarchy are sorted based on access frequencies of said resources.

9. A process as claimed in claim 1, including removing portions of said resources based on a metric of the relevance of said portions, prior to said step of generating cluster criteria.

10. A process as claimed in claim 1, wherein said step of generating said taxonomy of resource clusters includes iterative clustering of an existing cluster to generate sub-clusters of said cluster.

11. A process as claimed in claim 1, wherein said step of generating said taxonomy includes adding a resource to an existing cluster if the similarity of said resource to said cluster meets a similarity requirement.

12. A process as claimed in claim 1, wherein said step of generating cluster criteria includes generating a new cluster if the similarity of said resource to each existing cluster does not meet a similarity requirement.

13. A process as claimed in claim 1, wherein said step of generating cluster criteria includes selecting a similarity value for clustering on the basis of goodness values for respective groupings of said resources generated for respective similarity values.

14. A process as claimed in claim 13, wherein the goodness value for each grouping is generated on the basis of similarity values for resources within the clusters of the grouping and differences between similarity centroids for the clusters of the grouping and a global centroid for said resources.

15. A process as claimed in claim 1, wherein said steps of generating cluster criteria and generating a hierarchy of resource clusters are scalable with the number of said resources.

16. A process as claimed in claim 1, wherein said cluster criteria are generated from a subset of said resources.

17. A process as claimed in claim 16, including selecting said subset by random sampling of said resources.

18. A process as claimed in claim 16, including classifying resources into said resource clusters.

19. A process as claimed in claim 16, wherein said step of generating said taxonomy includes using clusters generated on the basis of said cluster criteria as a training set for a classifier.

20. A process as claimed in claim 19, wherein said classifier includes a support vector machine

21. A process as claimed in claim 16, wherein said step of generating said taxonomy includes clustering using a similarity value determined whilst clustering the subset of said resources.

22. A process as claimed in any one of claims 16 to 21, including generating one or more new clusters for resources that are not substantially similar to said resource clusters.

23. A process as claimed in claim 16, including maintaining the coherence of said clusters as the number of said resources increases by reducing said similarity value with increasing number of said resources.

24. A process as claimed in claim 16, including selecting a subset of clustered resources as the number of clustered resources increases to generate a metric for evaluating the quality of existing clusters.

25. A process as claimed in claim 24, including determining a new similarity value for reclustering said existing clusters on the basis of the quality of said existing clusters.

26. An information resource taxonomy system having components for executing the steps of any one of claims 1 to 25.

27. A computer-readable storage medium, having stored thereon program code for executing the steps of any one of claims 1 to 25.

28. An information resource taxonomy system, including

a data collector for collecting information resources from a communications network; and

a taxonomy generator for generating a taxonomy represented by a hierarchy of resource clusters, using cluster criteria generated from said resources.

29. A system as claimed in claim 28, including an editor for editing said criteria, and a renderer for generating linked document data for displaying said hierarchy.

30. A system as claimed in claim 28, including a classifier for classifying further resources collected by said system.

31. A system as claimed in claim 28, wherein said system is scalable with respect to the number of said resources.

32. A system as claimed in claim 28, including a parallel cluster search system for evaluating clusters in parallel.

33. A system as claimed in claim 28, including a parallel classifier for classifying further resources in parallel.

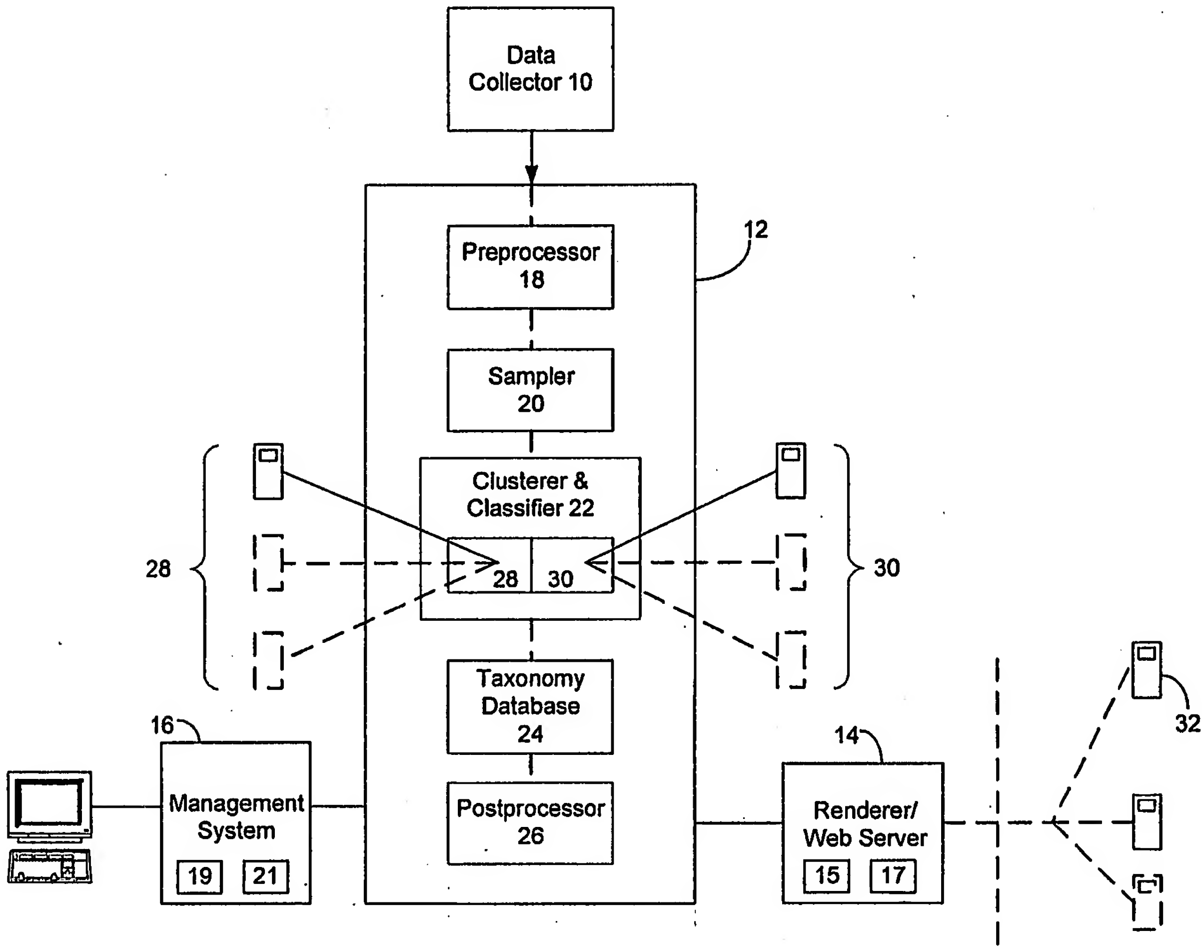

Data
Collector 10
Preprocessor
18
Sampler
20
Clusterer &
Classifier 22
28
30
Taxonomy
Database
24
Postprocessor
26
12
28
30
16
Management
System
19
21
14
Renderer/
Web Server
15
17
32

Figure 1

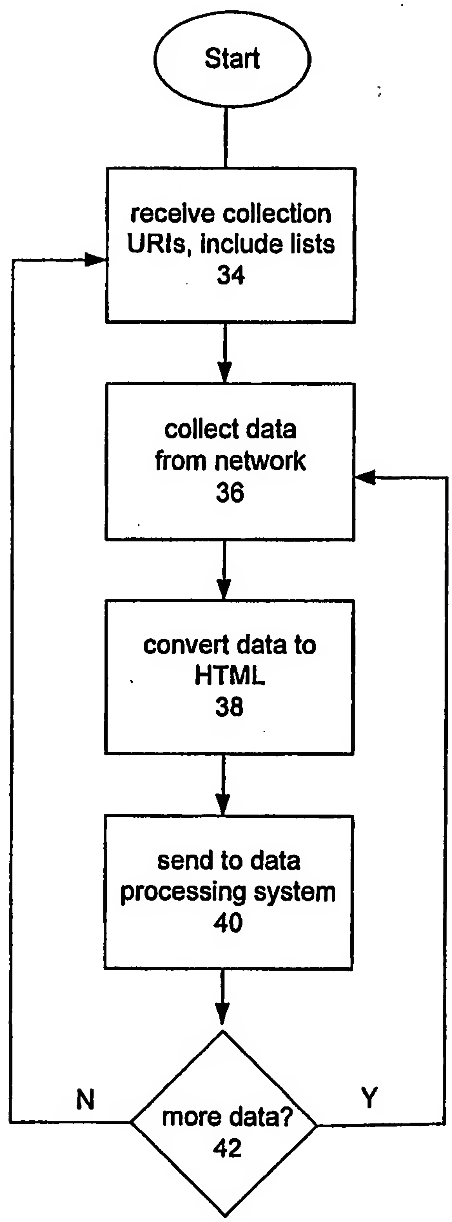

Start
receive collection URIs, include lists 34
collect data from network 36
convert data to HTML 38
send to data processing system 40
more data? 42
N
Y

Figure 2

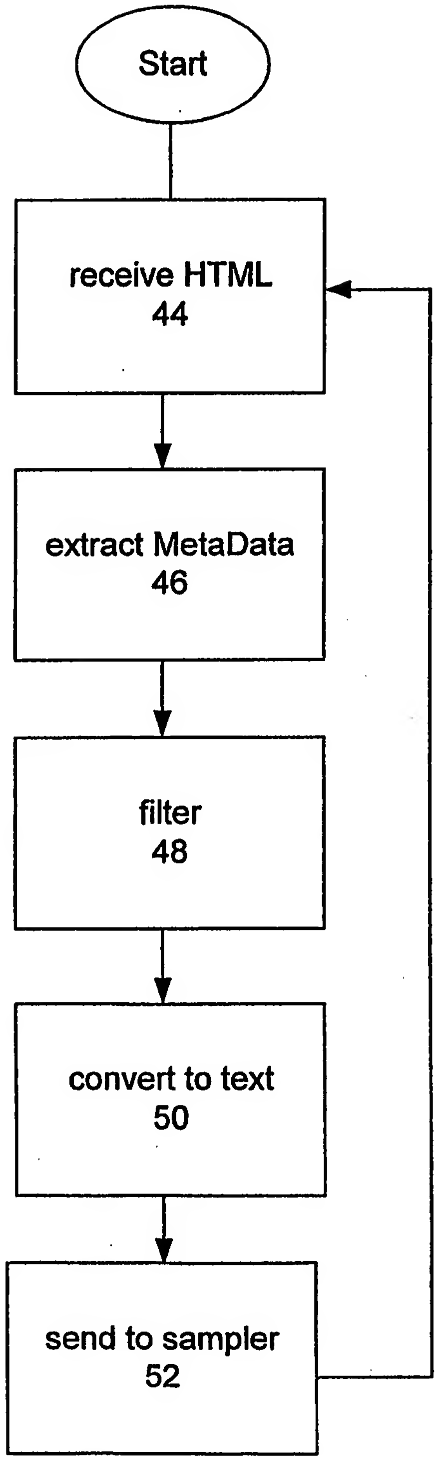

Start
receive HTML
44
extract MetaData
46
filter
48
convert to text
50
send to sampler
52

Figure 3

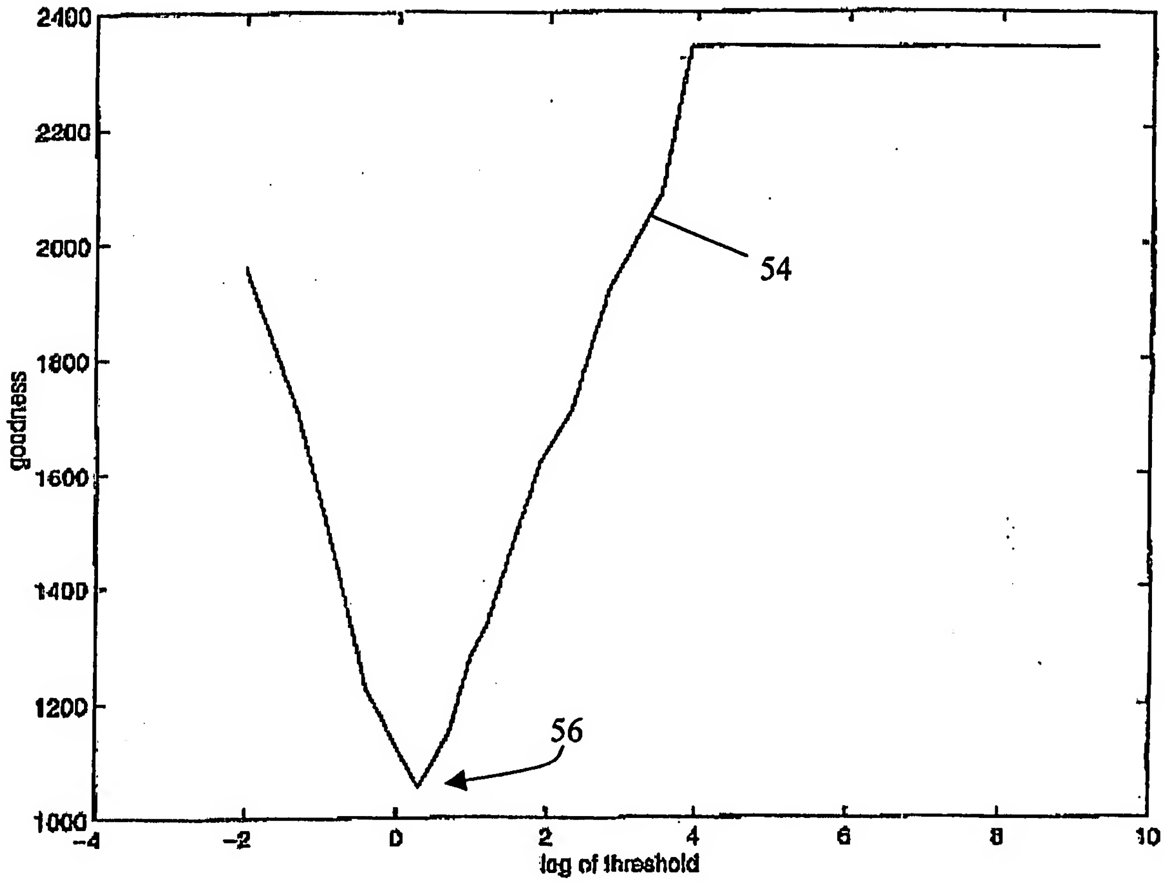

2400
2200
2000
1800
1600
1400
1200
1000
goodness
-4
-2
0
2
4
6
8
10
log of threshold
54
56

Figure 4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU02/01719

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl. [7]: G06F 17/30, 17/20

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
USPTO, DWPI (taxonomy, classification, categorisation, cluster, hierarchy)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 00/62203 A1 (SEMIO CORPORATION), 19th October 2000 the whole document	1-33
X	EP 0 704 810 A1 (HITACHI, LTD.), 3rd April 1996 the whole document	1-33
X	US 5,974,412 A (HAZLEHURST et al), 26th October 1999 the whole document	1-33

[X] Further documents are listed in the continuation of Box C [X] See patent family annex

* Special categories of cited documents:
- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier application or patent but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed
- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- "&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
14 February 2003	13 MAR 2003

Name and mailing address of the ISA/AU	Authorized officer
AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA E-mail address: pct@ipaustralia.gov.au Facsimile No. (02) 6285 3929	M. J. O'ROURKE Telephone No : (02) 6283 2017

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X, P	US 6,360,227 B1 (AGGARWAL et al), 19th March 2002 the whole document	1-33
X, P	US 6,446,061 B1 (DOERRA et al), 3rd September 2002 the whole document	1-33

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.
PCT/AU02/01719

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report		Patent Family Member			
WO	200062203	AU	200042212	EP	1208464
EP	704810	JP	8153121	US	5832470
US	5974412	US	6289353		
US	6360227	NONE			
US	6446061	NONE			

END OF ANNEX